ANONYMOUS RECOMMENDATION TECHNIQUE

PRIOR APPLICATIONS

The present invention is filed as a continuation-in-part of US Patent Application No. 09/944,739, titled ANONYMOUS ACQUISITION OF DIGITAL PRODUCTS BASED ON SECRET SPLITTING, filed on August 31, 2001 in behalf of Vora et al., the same inventors as in the present application, which prior application is assigned to the Hewlett-Packard Company, the same assignee as in the present application.

FIELD OF THE INVENTION

The present invention relates generally to electronic commerce systems and, in particular, to a technique for providing anonymous recommendations within such electronic commerce systems.

BACKGROUND OF THE INVENTION

Electronic commerce is increasingly becoming a part of everyday life. In particular, the rapid growth of the Internet and World Wide Web has lead to a corresponding increase in the ability to acquire goods and services remotely. A generalized example in accordance with current techniques is illustrated in FIG. 1. In particular, an entity 102, such as an individual or organization, may communicate with a provider 104 via a public network 103. The entity 102 transmits a variety of information to the provider 104 in order to acquire a product being offered by the provider 104. The information sent by the entity 102 typically comprises an

identification of the entity, an identification of the product being acquired and, optionally, information regarding the price of the product being acquired. In turn, where the acquisition is a purchase, the provider 104 may supply some or all of the information from the entity 102 to a credit agency 106. As a result, the provider 104 has specific knowledge of the products being purchased by the entity 102. Likewise, the credit agency 106 has specific knowledge that the entity 102 is purchasing products from the provider 104.

Based on the knowledge of what products or services are being acquired, the provider 104 can often supply one or more recommendations to the entity. That is, if an entity purchases, for example, tickets to a classical music concert, the provider may be able to recommend other classical music concerts or even other products, such as recordings of classical music. As providers become more familiar with additional specific activities performed by specific entities within the electronic commerce system (i.e., develop profiles about the entities), they can further refine their recommendations so as to provide more targeted or relevant recommendations with increased likelihood that the specific entities will follow up on the recommendations. While specific knowledge of a given entity's on-line activities may be valuable to the provider as a source of providing recommendations, which recommendation may even be helpful to or welcomed by some of the targeted entities, an increasing number of consumers object to commercial enterprises and other third parties having specific knowledge of their on-line activities. Clearly, a consumer's desire for privacy conflicts with the desire of commercial enterprises' to be able to recommend additional services and products based on past activities by consumers.

Therefore, a need exists for techniques that allow the formulation of recommendations based on some degree of knowledge about a given entity's on-line activities, but that also provides the given entity with a corresponding degree of privacy.

SUMMARY OF THE INVENTION

8. A commerce system comprises a plurality of entities each having an associated entity identity that is stored as a plurality of secret shares amongst at least a portion of a plurality of shareholders. Each plurality of secret shares comprises a

subset of a finite set of secret share values. An apparatus and method for the commerce system includes an association, for each entity of the plurality of entities, for at least one activity conducted by the entity within the commerce system with each of the plurality of secret shares used to store the entity identity corresponding to the entity such that each secret share of the finite set of secret share values has associated therewith a set of activities from at least a portion of the plurality of entities. A receiver for receiving sets of activities associated with each secret share of a first plurality of secret shares used to store a first entity identity corresponding to a first entity is included with a generator, coupled to the receiver, for generating an estimated activities list, for the first entity, comprising an intersection of the sets of activities.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram illustrating a typical arrangement used in electronic commerce in accordance with prior art techniques.

FIG. 2 is a block diagram illustrating an arrangement that may be used for electronic commerce in accordance with the present invention.

FIG. 3 is a flow chart illustrating a technique in accordance with the present invention.

FIGS. 4-18 illustrate an example of providing an estimated activities list in accordance with the present invention.

FIGS. 19-33 illustrate another example of providing an estimated activities list in accordance with the present invention.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

The present invention provides a technique for making recommendations to entities in an electronic commerce system (based, for example, on the Internet or World Wide Web) while simultaneously maintaining activities profiles of individual entities in secret. In the context of the present invention, activities encompass substantially all actions in which an entity provides its entity identification to a third party including, but not limited to, purchasing a digital object or service, providing information to a third party for the purposes, for example, of rating something in a

survey, etc. Furthermore, electronic commerce systems as used herein are not limited to systems supporting commercial transactions, but instead encompass all systems whereby an entity at least provides its entity identification to a third party for any purpose.

The present invention employs secret sharing techniques whereby information regarding particular activities by an entity is kept confidential and yet accessible when required to generate recommendations for the entity's benefit. An entity engaging in an activity within the electronic commerce systems supplies data such as an entity identification to an anonymity service. In turn, the anonymity service splits the entity identification into a plurality of secret shares that are thereafter provided to a corresponding plurality of shareholders. The nature of the secret splitting process is such that each shareholder is unable to reproduce the secret corresponding to the shareholder's share without the other shareholders involved in the process. Furthermore, the process whereby the secret shares are generated should be reproducible by the anonymity service, such that the anonymity service can generate all the sets of the plurality of secret shares ever generated for a particular entity identification. Information regarding the specific activity (e.g., information identifying the specific activity) is associated with each secret share of the plurality of secret shares used to memorialize the entity's identity. Assuming that there are a finite number of secret shares available from which any given entity's plurality of secret shares may be selected, over time and multiple activities conducted by multiple entities, each secret share will have associated therewith a set of activities corresponding to a plurality of entities.

When it is desired to generate recommendations for a given entity based on that entity's past activities, an intersection of the sets of activities associated with that entity's secret shares is provided as an estimated activities list. That is, activities found in each set of activities associated with each secret share of the given entity's secret shares are included in the estimated activities list, whereas activities found in less than all of the sets of activities associated with the entity's secret shares are excluded. Thereafter, one or more recommendations may be made based on the estimated activities list. By associating activities with a plurality of secret shares, the

present invention facilitates anonymous transactions, particularly anonymous recommendations.

The present invention may be more readily described with reference to FIGS. 2-33. Referring now to FIG. 2, there is illustrated a block diagram of a system 200 in accordance with the present invention. In particular, an anonymity service 203 is provided as an intermediary between entities 202 (e.g., acquirers of services or digital products) and one or more providers 204. In practice, the anonymity service 203 and provider 204 can be in communication with a clearing house and a credit agency in support of commercial transactions; for clarity, the clearing house and credit agency are not illustrated in FIG. 2. Although direct connections are illustrated between the anonymity service 203 and the various other elements of the system 200, it is understood that these connections may comprise paths established through public networks such as the Internet or World Wide Web, within private networks or through a combination of public and private networks.

In the context of the present invention, each of a plurality of entities 202 may comprise any individual or organization capable of conducting activities within the electronic commerce system 200. Such activities may comprise virtually any actions in which an entity is willing to include its entity identification. For example, such activities may include, but are not limited to, providing or acquiring something of value to/from a third party, such as participating in a survey, subscribing to a newsletter or affinity group chat room, or acquiring services or digital products from the provider 204. In the context of the present invention, digital products comprise anything capable of delivery via a communication network. For example, digital products may include downloadable software or digital data such as text, audio, video or images. Those having ordinary skill in the art will recognize that other types of digital products may be used in conjunction with the present invention, and the present invention is not limited in this regard. Regardless, each activity conducted within the system 200 should be susceptible to unique identification. For example, activities may include actions such as, "purchased video disk of 'Cinderella'", "downloaded trial version of XYZ Co.'s video game software", "donated money to Planned Parenthood", etc. In sum, activities in the context of the present invention

comprise any actions having associated therewith an entity identification and that are susceptible to a unique identification.

In practice, each entity 202 communicates with the anonymity service 203 via a computer implementing a network communication program, such as a browser or the like. The provider 204, in turn, may likewise comprise any individual or organization that provides services or digital products or that accepts information from entities via a communication network. More generally, the provider 204 may comprise any individual or organization that is the intended target of an entity's action(s) and is therefore a recipient of that entity's identification. The anonymity service 203 preferably comprises a computer-implemented service available via a communication network such as the Internet or World Wide Web. As depicted in FIG. 2, the anonymity service 203 preferably comprises a processor 210 and memory 212. For example, the anonymity service may be implemented using one or more network servers executing stored software routines as known in the art.

The anonymity service 203 is in communication with a plurality of shareholders 207. As described in greater detail below, each of the shareholders 207 is provided with a secret share (preferably representative of at least an entity's identification) which, by itself, does not enable an individual shareholder to reconstruct a secret, i.e., an entity's identity. In a preferred embodiment, the process used to generate the secret shares is reproducible; that is, for any given input, a predictable set of secret shares will be provided as output. Furthermore, it is preferred that the process used to generate the secret shares possess the property that any secret share is equally likely to take on a particular value as any other value. For example, if secret shares are expressed as n-bit binary words, there exists a finite set of secret share values comprising 2^n possible values. Thus, the likelihood of any one secret share value of the finite set of secret share values being generated is preferably equivalent to $\frac{1}{2^n}$.

The set of shareholders 207 is preferably static relative to the anonymity service 203 and to the number of secret share values in the finite set of secret share values. That is, for a given entity's identification, information corresponding to a given secret share value is always sent to the same shareholder of the set of

shareholders 207, which set itself is unchanging from the anonymity service's perspective. For example, if there are ten possible secret shares (S_1 through S_{10}) and ten possible shareholders (H_1 through H_{10}), a possible arrangement is to uniquely assign one of the ten secret share values to a corresponding one of the ten shareholders. Further, it is preferable to ensure that the first shareholder always receives the first secret share value produced, the second shareholder always receives the second secret share value produced, and so on. Of course, other arrangements are possible. If multiple anonymity services are provided, each could be associated with its own, well-described and not necessarily static, plurality of shareholders, which plurality of shareholders may or may not include shareholders that are also associated with other anonymity services.

In any event, each shareholder is capable of receiving secret shares and information regarding activities from the anonymity service 203. To this end, each shareholder preferably comprises a computer-implemented device capable of communicating with the anonymity service 203. Because secret sharing schemes are vulnerable to the extent that separate shareholders could collaborate to ascertain the secret in their possession, it is advantageous to maintain the identity of each shareholder in confidence from the other shareholders. Furthermore, it is preferred to select the shareholders such that they have an inherent reason not to collaborate with each other. For example, shareholders in possession of the secret shares corresponding to a single secret may comprise competitors in a given industry. Such competitors are inherently unlikely or unwilling to share information with each other. Additionally, the shareholders may comprise a privacy organization that is dedicated to advocating privacy in electronic commerce, and therefore unlikely to collaborate with other shareholders. Further still, the entity 202 may comprise one of the shareholders, or the shareholders 207 may be known to the entity 202, such as family members or friends or an Internet mailing list constructed around a common interest.

Referring now to FIG. 3, a method in accordance with the present invention is illustrated. In particular, the method of FIG. 3 is preferably implemented by the anonymity service 203. As shown, two parallel, independent paths are provided. On the left, a path is provided whereby activities by separate entities are added to sets of activities associated with separate secret shares. On the right, a path is provided

whereby a set of recommendations, based on the sets of activities associated with the separate secret shares, may be generated.

At block 302, it is continuously determined whether a given entity has transmitted its entity identification and information regarding one or more activities to the anonymity service. In an embodiment of the present invention, the entity identification and information regarding the one or more activities may comprise part of a service request sent to the anonymity service. In this manner, the anonymity service acts as an intermediary between the entity and the third party to which the entity's action activity is directed. The anonymity service can be provided as an additional service by an entity's on-line service provider, such as an Internet Service Provider (ISP). In one embodiment of the present invention, the information provided at block 302 is securely transmitted to the anonymity service. Security in the transmission of such information may be provided using known techniques, such as encryption or a trusted path.

The entity identification may comprise any unique identifier such as a public key, credit card number or the like. Likewise, the information regarding the one or more activities may comprise any unique value or description sufficient to distinguish one activity from another. In practice, at least with initial implementations, it is expected that such activity-describing information will be application dependent. For example, if the present invention is limited in its application to the realm of activities related to books (e.g., on-line purchases, ratings, etc.), the existing system of International Standard Book Number (ISBN) codes could be used to uniquely identify particular books referenced in a given activity. Similar use could be made of stock keeping unit (SKU) codes or similar codes currently provided on many products. More generally, various schemes have been proposed, such as the so-called Consumer Profile Exchange Standard (CPEX) and Digital Object Identifiers (DOI) that may be applicable across a broad spectrum of activities, and may therefore be beneficially applied to the present invention.

When an entity has transmitted its entity identification and information regarding the one or more activities, processing continues at block 304 where a cryptographic or other secret splitting technique is used to split the entity identification into a plurality of secret shares. Such secret splitting techniques are

well known in the art. In essence, a secret splitting technique takes a secret and divides it up into pieces such that each piece by itself does not allow a holder of that piece to reconstruct the secret. However, a holder in possession of all of the pieces is able to reconstruct the secret.

The present invention requires well-described and reproducible secret splitting, i.e., a secret splitting technique that results in reproducible sets of secret share values for a given entity identification. A number of cryptographic secret sharing schemes use random number generation and, as a result, the secret share values are generally not reproducible. Typically, in such schemes, a seed value is typically used to initialize a process that generates a stream of essentially random data, which random data (i.e., the secret shares) is then used to secure other data. Random number generation usually enhances the secrecy of a scheme because a well-defined, predetermined relationship does not exist between the secret share values and the secret, and the shares hence reveal less information about the secret. As a practical matter, however, the present invention cannot be implemented using a secret splitting/sharing scheme that results in secret shares that are different each time for identical inputs. Rather, the present invention is preferably implemented using a secret splitting scheme that generates one of a finite number of possible sets of secret shares each time for identical inputs. The number of possible secret share sets for a given entity identification should be small compared to: (a) the average number of times an entity is expected to use the service, and (b) the total number of sets of secret shares possible (for m n-bit secret shares, this number is 2^{nm}). Further, depending on usage and number of activities, the size of the number of possible secret share sets for any given entity identity also affects recommendation accuracy.

Most of the cryptographic secret sharing schemes that require random numbers can be easily modified for use with the present invention. For example, in one approach, random numbers can be generated once, stored with an entity identification by the anonymity service, and used every time the entity identification is received thereafter. In another approach, the same random number can be used for all entity identities.

However, both of these approaches reduce the secrecy of the schemes. The first approach provides more privacy to the consumer relative to the shareholders

because the shareholders have data from only one entity to reverse-engineer in order to obtain information about the corresponding random number and, through it and a single secret share value, the entity identification. However, because the first approach requires storage of entity identifications by the anonymity service, (not required by the second approach), it provides severely reduced privacy to the consumer with respect to the anonymity service.

The second approach provides less privacy to the consumer relative to the shareholders because they might be able to analyze patterns of the shares corresponding to different entity identifications to make educated guesses about the single random number, and from that information also determine information about entity identification from a single secret share value. Similar schemes, with the fixed numbers changing every so often, even at random, and being stored against entity identification for later reference, can also be used to provide more privacy to the consumer relative to the shareholders, though these will still be privacy-compromised to some degree with respect to the anonymity service and could decrease recommendation accuracy.

A third approach is for the anonymity service to use a function of the entity identification to generate the random data for that entity identification so that the random data does not have to be stored. The service can thereafter construct the random data each time it is needed. Because each entity identification is presumably unique, the resulting random data will not be the same for all entity identifications, and the function used can be changed occasionally if required. In particular, the function can be randomly chosen from a small, finite set of functions. The scheme could be made more secure if the function were a so-called one-way function, i.e., a function that is easy to compute but difficult to invert. This would make it harder to recover an entity identification from a secret share value. Those having ordinary skill in the art will recognize that other types of permutations may be used to generate reproducible shares in conjunction with the present invention, and the present invention is not limited in this regard.

An exemplary secret sharing scheme that uses random numbers, and its modification for use with the present invention, is now described. As an example of secret sharing, assume that a party A wishes to split a secret S into three shares that

will be subsequently given to parties B, C and D. In accordance with a preferred embodiment of the present invention, further assume that the secret S is represented as a string of bits having length M. First, A generates two random bit strings, X and Y, each of length M. (Techniques for generating random bit strings are well known in the art of cryptography and are therefore not described in detail herein.) The secret S is thereafter exclusive-OR'd with X and Y to provide a new bit string Z, also of length M:

$$Z = S \oplus X \oplus Y$$

Thereafter, A provides Z, X and Y (the secret shares) to, for example, B, C and D (the shareholders), respectively. Note that none of B, C or D is able to reconstruct the secret S based solely on their respective share (Z, X or Y). To the contrary, the only way to reconstruct the secret is to combine the secret shares once again:

$$S = Z \oplus X \oplus Y$$

The above-described scheme can be modified to generate reproducible secrets instead of random secrets. Instead of random bit strings, X and Y could be outputs of well-defined functions of the entity identification. For example, X and Y could be squares, cubes, and/or n^{th} powers of the entity identification. Because X and Y would no longer be random, the security of the secret splitting scheme would be reduced. However, this is a cost of obtaining X and Y values that are reproducible. To mitigate the impact on the security, the function can be changed, e.g., to the m^{th} power, after a certain period of time, or it can be a different function depending on parameters such as the day of the week, or it can be a function randomly chosen from a finite set of functions (for example, the function could be one of the $101\text{-}105^{th}$ powers). The total number of different functions used for each entity identification should be small compared to the average number of times each entity identification is used in the anonymity service, and small compared to the total number of sets of secret shares

possible. The larger the number of functions, the more privacy is protected from the secret shareholders at the possible expense of recommendation accuracy.

While this is a simple example, it illustrates the basic concept and implementation of secret splitting modified for reproducibility. Simple extensions will be evident to those having ordinary skill in the art. For example, a larger number of shareholders may be employed by simply generating additional reproducible bit strings to combine with the secret. One publication teaching a variety of cryptographic secret splitting techniques is "Applied Cryptography" by Bruce Schneier (John Marley & Sons, 1996), the teachings of which are incorporated herein by this reference. Referring back to FIG. 3, the number of secret shares provided at block 304 for the entity identification is a matter of design choice. However, in a preferred embodiment, the number of secret shares is always the same and the secret splitting technique is reproducible, though equivalent outputs may not always result from equivalent inputs.

So-called perfect secret splitting schemes are those schemes that result in secret shares in which any given secret share does not provide any information to an adverse party by which the secret may be identified. A number of known perfect secret sharing schemes result in at least one random secret. Modifying a perfect scheme to produce reproducible shares as described earlier could make the scheme non-perfect.

Besides cryptographic secret splitting schemes, other schemes which require less computation and in which individual secret shares do provide some information regarding the secret may also be used, although are not preferred. A simple illustration of a non-cryptographic secret splitting technique would be to split the secret up into constituent parts and providing those parts as the secret shares directly. For example, if a given entity's identification is simply "John Smith", the secret shares could be provided as the nine letters in the identification, i.e., "j", "o", "h", "n", "s", "m", "i", "t" and "h". While no one secret share (letter) allows a given shareholder to reconstruct the entire secret, it does provide some information about the secret.

Regardless, at block 306, the secret shares created at block 304 are sent to shareholders along with the information regarding the one or more activities. While

the secret shares could be sent to the shareholders in encrypted form in order to enhance security, the secret shares are sent unencrypted in a presently preferred embodiment. Likewise, the information regarding the one or more activities may be sent in an encrypted form or otherwise secure manner, but this is not preferred. Nevertheless, the information regarding the one or more activities is sent to the shareholders each time a secret share is produced.

The length of time required by each shareholder to store a corresponding secret share is a matter of design choice and may be dictated, for example, by legal requirements setting the length of time documentation regarding a transaction is to be stored. Likewise, the duration for which the information regarding the one or more activities is kept is a matter of design choice. In one embodiment of the present invention, the information regarding the one or more activities is kept indefinitely. In another embodiment, however, activities are assigned an expiration date such that they are no longer available after the expiration date. Further still, activities could be continuously inspected to see if no one has engaged in a certain activity for a certain period of time, in which case the activity is purged. Those having ordinary skill in the art will recognize that other schemes may be possible or desirable. Regardless, once a secret has been split and sent to the respective shareholders, the anonymity service discards any copies of the secret. In essence, the anonymity service consumes each secret and distributes the resulting secret shares to corresponding shareholders. An exception to this, albeit not a preferred one, arises where the anonymity service stores the random number associated with a particular entity identification, as described previously.

At block 308, the one or more activities sent to the shareholder are associated with the secret share. In a preferred embodiment, this is accomplished by each shareholder maintaining one or more profiles of activities in, for example, a computer-readable medium such as a digital memory device or the like. Each shareholder maintains a separate activity list for each secret share value. An example illustrating this is provided with reference to FIGS. 4 and 5.

FIG. 4 illustrates an exemplary system in which it is assumed that the finite set of secret share values comprises only ten values (S_1 through S_{10}). The number of possible secret share values in this example has been kept low for ease of illustration;

in practice, this number would be substantially larger. Additionally, it is assumed that there are only twenty possible activities (labeled "A" through "T"). Again, the number of possible activities in this example has been kept low for ease of illustration; in an actual implementation, this number would be significantly larger. It is further assumed that the identification of each entity (user) is split into five secret shares, which five secret shares are equally likely take on any of the ten possible secret share values. Note that the secret share values derived for any given secret may result in multiple occurrences of the same secret share value, e.g., the identification of User 1, when split, results in two instances of the S_5 secret share value. Finally, there are five shareholders available to receive, according to a predefined distribution scheme, the five secret shares generated for any given secret. For example, for each secret split into the five secret shares, a first secret share is provided to a first shareholder, a second secret share is provided to a second shareholder and so on. In a preferred embodiment, the activities for any given entity are stored in the same profiles to the extent that the secret splitting scheme always provides a reproducible output in response to the given entity's identity, which output causes the information regarding that entity's activities to always be sent to the same shareholders for association with the same secret share values. When the secret splitting scheme does not always provide the same output, but provides one of a finite set of outputs, each generated by the use of a different function as described earlier, the activities will go to different profiles. The example described herein, for ease of illustration, corresponds to the situation where the secret shares generated for a particular entity identity are always the same.

In FIG. 4, eight exemplary users are shown and their corresponding secret shares. The column numbers 1-5 above the secret share values correspond to the shareholders to which each secret share value is sent. That is, the first column of secret share values sets forth those secret share values sent to the first shareholder; the second column sets forth those secret share values sent to the second shareholder; etc. Additionally, each user is assumed to have taken part in three activities randomly chosen from the twenty possible activities. The limit of three activities was chosen for ease of illustration. In practice, the number of possible activities engaged in by a given entity could be less and, in many instances, would likely be more. Note that

some activities may considered to be more popular, e.g., activity "J", in that multiple users have engaged in those activities. Regardless, profiles associated with each secret share value sent to the first shareholder (Shareholder 1) in this example are further illustrated in FIG. 5. Each profile is the result of the activities engaged in by users whose identity, when split according to the secret splitting scheme, results in one of the secret share values shown in FIG. 5. For example, note that the "J" activity is reflected in the each of the S_1, S_2 and S_3 profiles by virtue of it being included at least in the activities of Users 1, 2 and 3. Furthermore, even though an activity is included in the activities of multiple users, it is reflected in any given profile only once. Exemplary profiles maintained by each shareholder in the example based on FIG. 4 are shown in each of FIGS. 5-9.

Referring once again to FIG. 3, the process of generating recommendations for a given entity is illustrated by the path comprising blocks 310-316. At block 310, the sets of activities (e.g., the profiles) associated with each secret share corresponding to the given entity are obtained. For example, referring once again to FIGS. 4 and 5, if it was desired to generate recommendations for User 4, the profiles associated with secret shares S_1, S_3, S_5, S_7 and S_{10} would be obtained. Because each profile includes information regarding not only the activities of User 4, but a number of the other users as well, knowledge of any one profile does not allow an adverse party to reconstruct User 4's actual activities list. The process of block 310, in implementation, requires that the given entity's identification is first obtained and then split into a corresponding set of secret shares. This set of secret shares is then sent to the corresponding shareholders as part of requests to obtain the profiles corresponding to those secret shares. Once again, this is possible assuming that the secret splitting scheme is statically reproducible. With regard to the current example (FIG. 4), FIGS. 10-17 illustrate, for each user, the profiles corresponding to the secret share values generated by the user's identity.

At block 312, an estimated activities list corresponding to the given user is generated based on the profiles obtained at block 310. To this end, an intersection of the multiple profiles from block 310 is calculated. The estimated activities list thus comprises only those activities that are found in each of the profiles obtained at block 310. In one embodiment of the present invention, such estimated activities lists may

be stored on a computer-readable medium. With regard to the example based on FIG. 4, this is further illustrated with regard to FIGS. 10-18. Thus, in the estimated activities lists shown in FIG. 18, each activity is seen to occur in each profile associated with the given user. For example, with respect to User 1 (FIG. 10), the intersection of the activities associated with secret shares S_3, S_5, S_7, S_5 and S_{10} results in an estimated activities list comprising activities "D", "F", "G", "J", "L" and "M", as shown. Similar estimated activities lists for each of the other seven users (FIGS. 11-17), as shown in FIG. 18, are calculated in the same manner. In order to preserve the anonymity of the entities, the estimated activities lists are preferably generated apart from the corresponding entity identifications.

The nature of the process used to generate estimated activities lists is such that activities not actually engaged in by a given entity can be added to the estimated activities list for that entity. Such "added" activities are illustrated in boldfaced and italicized type in the estimated activities lists of FIG. 18. For example, activity "J", appearing in each profile, is frequently added to the estimated activities lists of entities that did not actually engage in activity "J". Additionally, activities "D" and "L", although not occurring in every profile, occur with sufficient frequency that they too are inaccurately added to the estimated activities lists of User 5. Such inaccuracies are a cost of being able to provide recommendations in an anonymous manner. The likelihood of such additions to the estimated activities lists is lowest when each activity is equally likely. Each activity is not equally likely in the example shown, where activity J is much more likely than other activities.

The effect of greater uniformity in distribution is shown in FIGS. 19-33. Comparison of FIG. 19 with FIG. 4 reveals that the occurrences of activity "J" for Users 1 and 4 have been replaced with occurrences of activity "K" and "F", respectively. The profiles resulting from the example illustrated in FIG. 19 are illustrated in FIGS. 20-24. Note that in the profiles shown in FIGS. 20-24, when compared to those in FIGS. 5-9, the activities are closer to being equally likely. While the example does not illustrate this, this could also slightly compromise entity privacy, which, as alluded to above, usually bears an inverse relationship with recommendation accuracy. Upon collecting the relevant profiles for each user, as illustrated in FIGS. 25-32, the intersections of the profiles for each user may be

calculated leading to the estimated activities list shown in FIG. 33. Note that the estimated activities lists illustrated in FIG. 33 are slightly more accurate than those illustrated in FIG. 18 in that the estimated activities for User 6 do not include an otherwise inaccurate activity "J".

When the secret shares are reproducible but not always identical for a given entity identity, the above described process must be repeated for each set of secret shares corresponding to the entity identity. This involves more computation and could reduce recommendation accuracy, but will improve entity privacy with respect to shareholders.

Referring once again to FIG. 3, at block 314, a set of recommendations may be generated based on the estimated activities list. In the context of the present invention, a recommendation may comprise any information regarding goods, services, opportunities or any other information that may spur the recipient of the information on to further activities. For example, if the estimated activities list indicates that the corresponding entity has recently applied for a mortgage using an on-line broker, a set of recommendations may comprise the names of various firms that provide homeowner's insurance. Those having ordinary skill in the art will recognize that a variety of techniques exist for establishing a set of recommendations based on knowledge of previous (estimated) activities. Thereafter, at block 316, the recommendations are provided to the corresponding entity using the entity's identification. Preferably, this is accomplished by reconstructing the entity identification based on the secret shares for that entity. To this end, the anonymity provider retains, for a given entity, the identifications of the shareholders in possession of secret shares needed to reconstruct the entity identification. Furthermore, the anonymity service may retain transaction identifications that identify the particular shareholders needed relative to a given transaction.

For example, assume an entity purchases goods from a provider via the anonymity service. A technique for providing anonymity in such a transaction is disclosed in co-pending U.S. Patent Application No. 09/944739 and titled ANONYMOUS ACQUISITION OF DIGITAL PRODUCTS BASED ON SECRET SPLITTING, the teachings of which application have been previously incorporated herein by reference. As disclosed in that application, entities are able to acquire

digital products in an anonymous fashion using secret splitting techniques such that third parties (including providers) are unable to identify specific acquisitions made by a given entity. Nevertheless, a provider may wish to make further recommendations for other products based on one or more transactions by a given entity. Using the presently-disclosed technique, the anonymity service can construct a set of recommendations and, using the techniques disclosed in the co-pending application, provide the recommendations in a manner that preserves entity privacy.

In the foregoing specification, the invention has been described with reference to specific embodiments. However, those of ordinary skill in the art will appreciate that various modifications and changes can be made without departing from the scope of the present invention as set forth in the claims below. Accordingly, the specification and figures are to be regarded in an illustrative rather than a restrictive sense, and all such modifications are intended to be included within the scope of present invention. For example, secure multi-party computing could be used in place of the anonymity service. That is, rather than a single third party managing anonymous transactions, a distributed model may be employed. As known in the art, secure multi-party computation involves computing a function of many variables where many parties each provide one or more variables. For example, secret shares of public keys may be used to encrypt and decrypt data without reconstructing the public key. As a result, a provider could send a set of recommendations to an entity in an encrypted form by letting the shareholders encrypt the recommendation using secure multi-party computation. Thus, in the context of the present invention, the shareholders themselves may implement functions described above as being implemented by the anonymity service (if the shareholders are known to each other) using known techniques.

Benefits, other advantages, and solutions to problems have been described above with regard to specific embodiments. However, the benefits, advantages, solutions to problems, and any element(s) that may cause any benefit, advantage, or solution to occur or become more pronounced are not to be construed as a critical, required, or essential features or elements of any or all the claims. As used herein, the terms "comprises," "comprising," or any other variation thereof, are intended to cover a non-exclusive inclusion, such that a process, method, article, or apparatus that

comprises a list of elements does not include only those elements but may include other elements not expressly listed or inherent to such process, method, article, or apparatus.